

CLAIMS

Sub 138
5 1. A method for extracting information from a database, which comprises a number of data tables containing values of a number of variables, each data table consisting of at least one data record including at least two of said values, said information being extracted by evaluation of at least one mathematical
10 function operating on one or more selected calculation variables, said extracted information being partitioned on one or more selected classification variables, characterised by the steps of:

15 identifying all data tables containing at least one value of one of said selected variables, such data tables being boundary tables;

20 identifying all data tables that, directly or indirectly, have variables in common with said boundary tables and connect the same, such data tables being connecting tables;

electing a starting table among said boundary and connecting tables;

25 building a conversion structure that links values of each selected variable in said boundary tables to corresponding values of one or more connecting variables in said starting table; and

30 evaluating said mathematical function for each data record of said starting table, by using said conversion structure to convert each value of each connecting variable into at least one value of at least one corresponding selected variable, such that said evaluation yields a final data structure containing a result of said mathematical function for every unique value of each classification variable.

35 2. A method as set forth in claim 1, characterised by the further step of presenting relevant

parts of said resulting data structure to the user in human-readable form.

3. A method as set forth in claim 1 or 2, characterised by the further step of initially
5 reading said data records of said database into the primary memory of a computer.

Sub A1-7
10 4. A method as set forth in any one of the preceding claims, characterised by the further step of initially assigning a different binary code to each unique value of each data variable in said database and storing the data records in binary-coded form.

15 5. A method as set forth in any one of the preceding claims, characterised by the further steps of initially identifying all data tables in said database that have variables in common, and assigning virtual connections between such data tables, thereby creating a database with a snowflake structure, wherein said connecting tables are located between said boundary tables in said snowflake structure.

20 6. A method as set forth in any one of the preceding claims, characterised by the further steps of identifying all calculation variables for which the number of occurrences of each value is necessary for correct evaluation of said mathematical function,
25 defining a subset of data tables consisting of boundary tables containing such variables and data tables connecting such boundary tables, electing said starting table from said subset, and including data on said number of occurrences of each value in said conversion
30 structure.

35 7. A method as set forth in any one of the preceding claims, characterised in that said starting table is the data table among said boundary and connecting tables having the largest number of data records.

8. A method as set forth in any one of the preceding claims, characterised by the further step of

Sub A17
5 building said final data structure, which includes a number of data records, each of which contains a field for each selected classification variable and an aggregation field for said mathematical function, wherein said building step includes sequentially reading a data record of said starting table, creating a current combination of values of said selected variables by using said conversion structure to convert each value of each connecting variable in said data record into a value of
10 at least one corresponding selected variable, evaluating said mathematical function for said current combination of values, and aggregating the result of said evaluation in the appropriate aggregation field based on the current value of each selected classification variable.

15 9. A method as set forth in any one of claims 1-7, characterised by the further step of creating a virtual data record containing a combination of values of said selected variables, wherein said creating step includes reading a data record of said starting table and
20 using said conversion structure to convert each value of each connecting variable in said data record into a value of at least one corresponding selected variable.

25 10. A method as set forth in claim 9, characterised by the further step of building said final data structure which includes a number of data records, each of which contains a field for each selected classification variable and an aggregation field for said mathematical function, wherein said building step includes sequentially reading a data record of said
30 starting table, updating the content of said virtual data record based on the content of each such data record, evaluating said mathematical function based on said updated virtual data record, and aggregating the result of said evaluation in the appropriate aggregation field
35 based on the current value of each selected classification variable in said updated virtual data record.

11. A method as set forth in claim 9, characterised by the further step of building an intermediate data structure which includes a number of data records, each of which contains a field for each
5 selected classification variable and an aggregation field for each mathematical expression implied by said mathematical function, wherein said building step includes sequentially reading a data record of said starting table, updating the content of said virtual data
10 record based on the content of each such data record, evaluating each mathematical expression based on said updated virtual data record, and aggregating the result of said evaluation in an appropriate aggregation field based on the current value of each selected
15 classification variable in said updated virtual data record.

12. A method as set forth in claim 11, characterised in that said step of building said intermediate data structure includes:
20 eliminating one of said classification variables in said intermediate data structure by aggregating said results over all values of said one classification variable for each unique combination of values of remaining classification variables, by creating
25 additional data records, and by incorporating said aggregated results in said additional data records of said intermediate data structure.

13. A method as set forth in claim 11 or 12, characterised by the further step of evaluating said
30 mathematical function based on said results in said aggregation fields for each unique combination of values of said classification variables, thereby building said final data structure.

Sub 7
35 14. A method as set forth in any one of the preceding claims, characterised in that said step of building said conversion structure includes:

Sub-A17

a) reading data records of a boundary table, and creating a conversion structure including a link between each unique value of at least one connecting variable in said boundary table and each corresponding value of at least one selected variable therein;

b) moving from said boundary table towards said starting table;

c) if a connecting table is found, reading data records of said connecting table, and substituting each unique value of said at least one connecting variable in said conversion structure for at least one corresponding unique value of at least one connecting variable in said connecting table; and

d) repeating steps (b)-(c) until said starting table is found.

15. An article of manufacture comprising a computer-readable medium having stored thereon a computer program for effecting the steps of a method for extracting information from a database as set forth in any one of the preceding claims.